Introduction to AI（CS103）– 05

The Basis of AI - Artificial Neurons

Jimmy Liu 刘江

2023-10-20

Lecture 5

Course Project/Survey

组号	项目类型	组长	组员	主题
1	project	虞快	杨家鉴、徐璟源、黄增荣	AI 诗人
2	project	陶文晖	郑青禾，邱天润，赵伟栋，刘天恩、李玉金	数字化记忆灾害
3	project	朱家润	冯星洋、刘乐平、范书豪	面向环扫影像的多结构分隔智能系统研发
4	project	陈旭阳	刘泽梁、赵俊儒、李家成	emotion recognition algorithm for EEG
5	project	谢毅	谢毅、李昭毅、吴鎏亿、王梓硕	垃圾邮件分类器
6	project	李岩	黄少霖、杨泽扬、朱丹、王奕鸣	基于AI技术的五子棋状态评估及自适应难度人机对弈
7	survey	杨祎勃	李麒飞、陈鹏如、杨一轩、张阳、钟洋	调研各类演化算法并进行归纳和分析
8	survey	吴成钢	吴成钢、潘乐宇、柏蔚泽、邢宇珩	智能交通
9	survey	曾宪清	陈文雁，秦颢轩，陈睿瑶，唐培致，胡清畅	人工智能与自然语言处理

Course Project/Survey

组号	项目类型	组长	组员	主题
10	survey	蒋浩天	冯海波，殷子尧，秦李旸，边政赫，张锦泽	人工智能在医疗适老化中的应用
11	project	李博洋	张羽乐、申家豪、徐皓鑫，徐龙翔，马怀远	2D超声骨面重建
12	project	冯泽欣	李子豪，黄子勖，张伟祎，陈奕冲，段柱材	SSL自监督学习在OCT图像上的运用
13	project	黄宇航	廖泽通、李明、赵可泰、罗嘉俊	人工智能方法在量化金融领域的应用
14	survey	罗雨涵	罗雨涵 余子越 周拓	少标签情况下的医疗图像分割
15	project	王德涵	杨博桥 郑荣菲 冯泉骊 王子铭	图书馆自画像
16	project	杨烜	刘圣鼎，张展玮，勾业备	基于AI技术的南科大讲座海报著录与分析
17	survey	杨宗奇	杨浩庭，张乐之，冯秋皓，宋一鸣，何昊东	人工智能在智能家居领域的应用综述

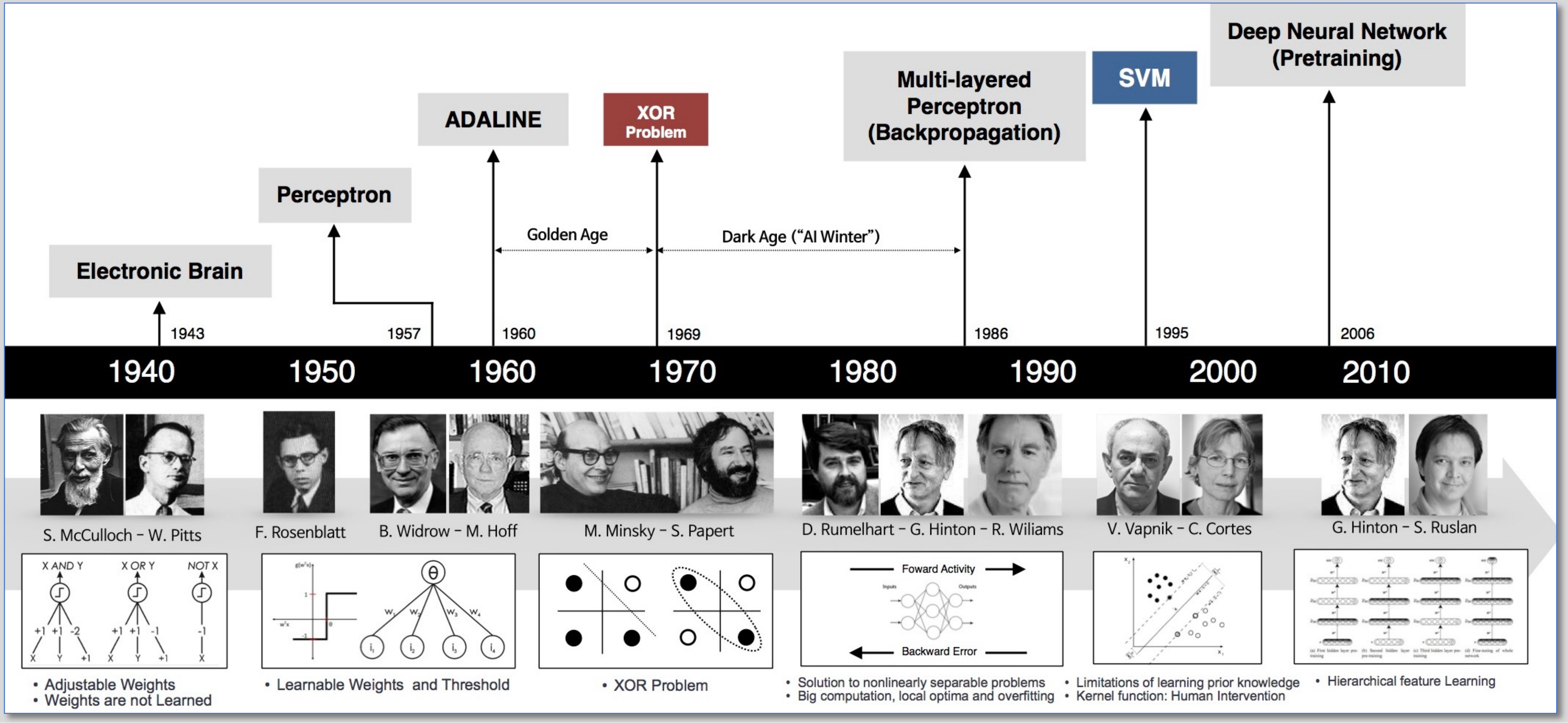
Course Project/Survey

组号	项目类型	组长	组员	主题
18	project	徐进哲	王梓鑫，盛鹏，刘亦辰	TBD
19	survey	李轲	李轲，肖李诚，李国豪，金仁良，张维宁	人工智能聊天
20	survey	TANG RYAN TZE HOU	ANTHONY BRYAN, FARIDA FITRIA ZUSNI, NGUYEN THANH LAM, CLAUDINE MARIETTA	Artificial Intelligence in Graphic Rendering
21	survey	陈茜	陈蕾，刘思远，廖子良，王浩然，王宇琛	人工智能在自然语言处理领域的综述

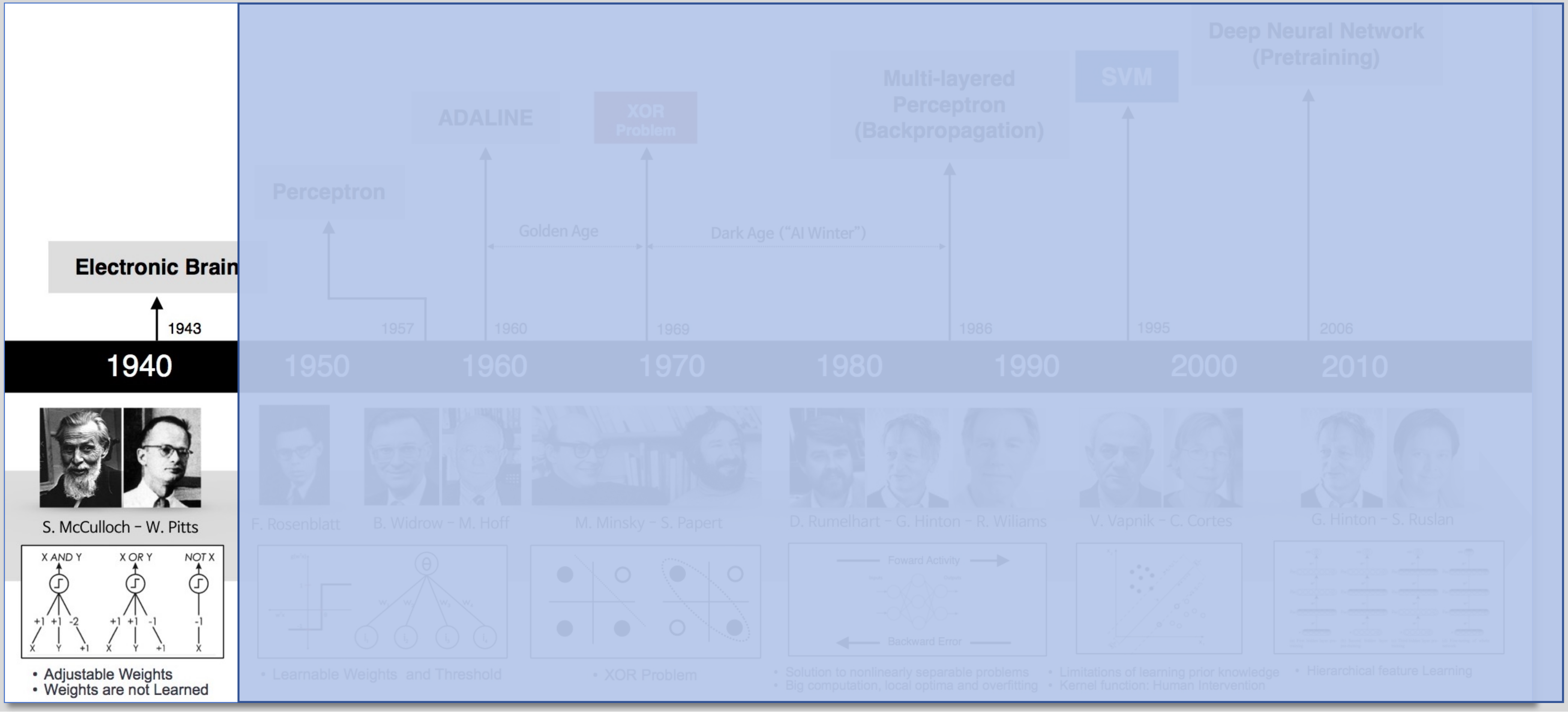
Lecture 5

Q1:List Some Characteristics of Biological Neurons

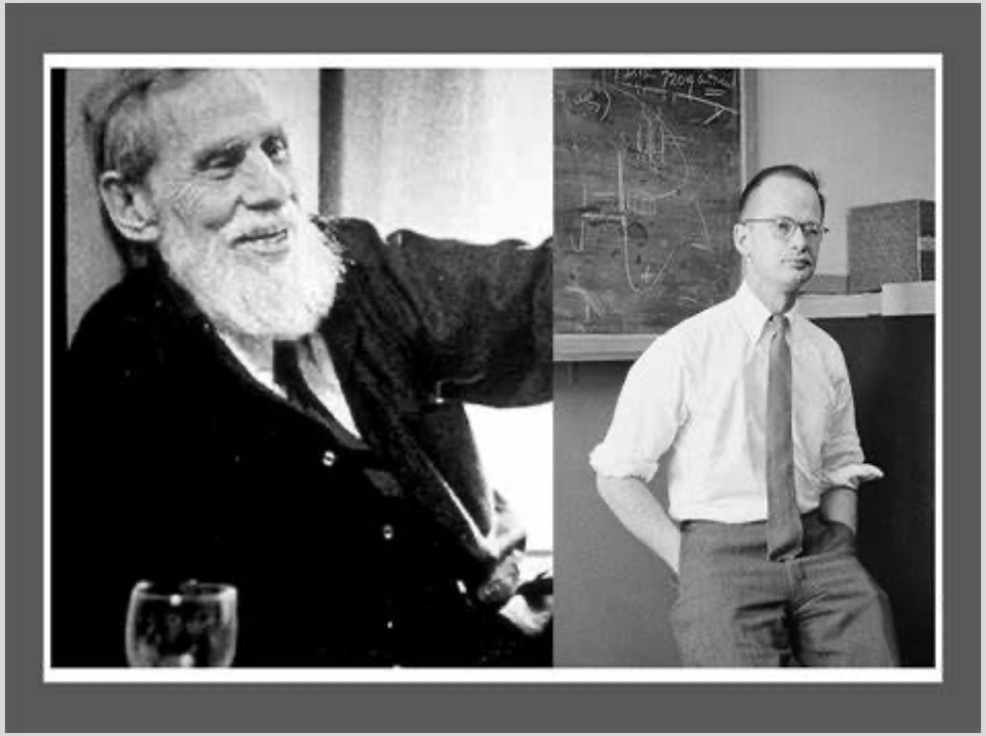
AI Algorithm Development Milestones

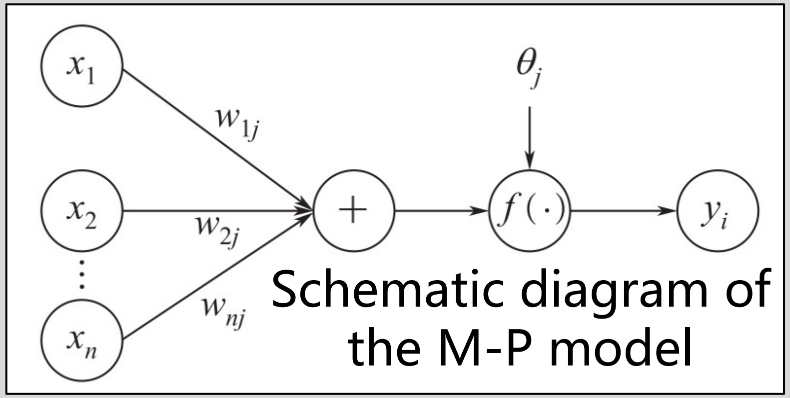


AI Development- Electronic Brain



M-P Model

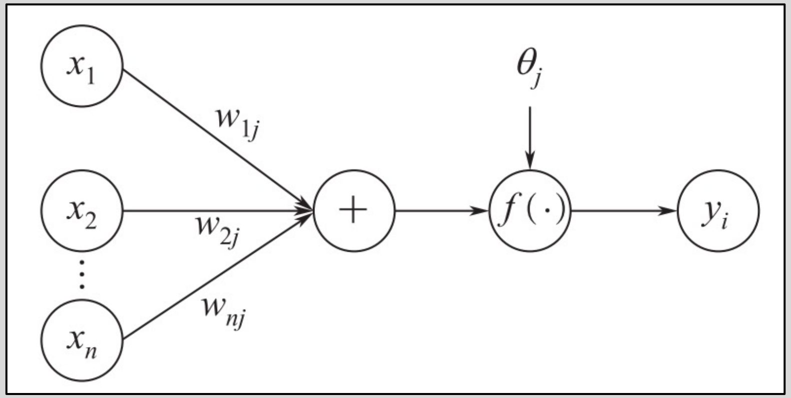




Schematic diagram of the M-P model

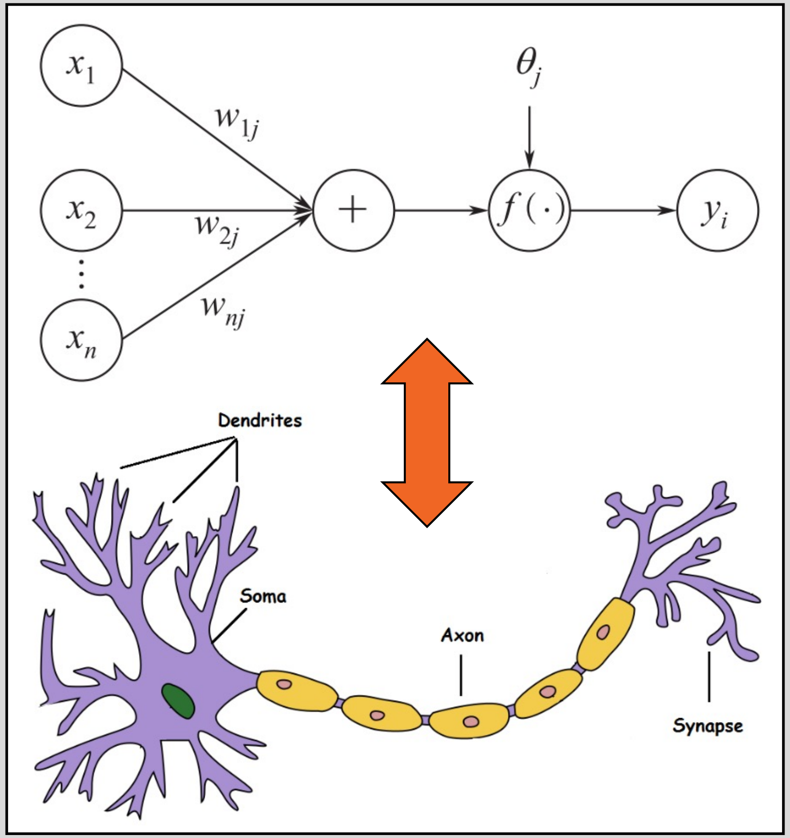
The **McCulloch-Pitts (M-P) model** of the artificial neuron was first proposed by the American psychologist W. McCulloch and mathematician W. Pitts in 1943 in their paper, "A logical calculus of the ideas immanent in nervous activity," based on the characteristics of biological neurons.

M-P Model
-Artificial Neuron Working Mechanisms



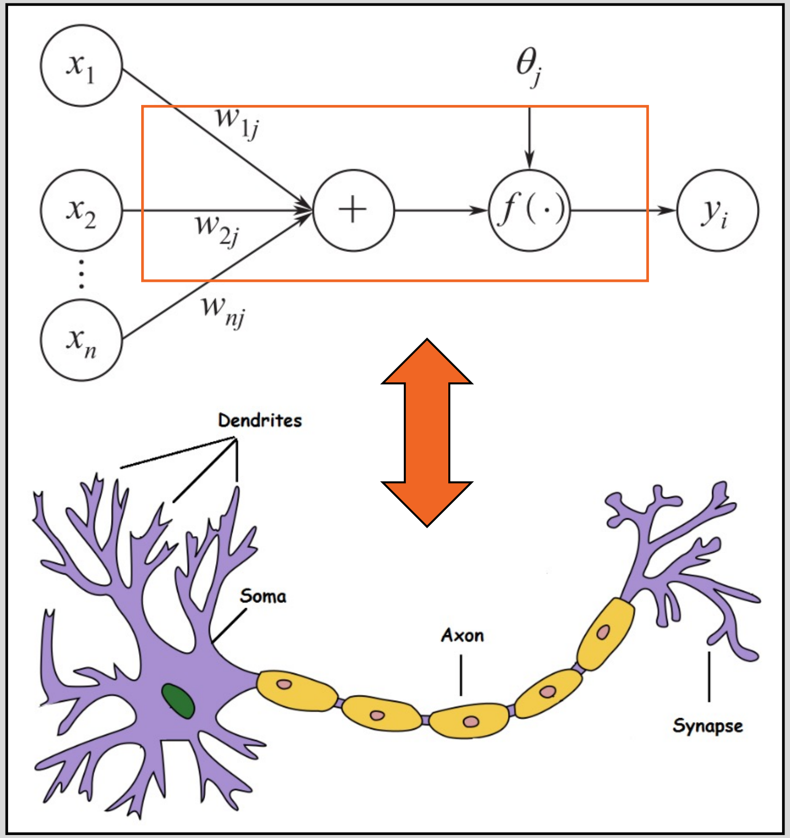
- The M-P model is a **multiple-input single-output** information processing unit;
- The inputs of artificial neurons can be categorized into two types: **excitatory inputs and inhibitory inputs**;
- That artificial neurons have **spatial integration properties** and **threshold properties**;
- There is a **fixed time delay between the input and output** of an artificial neuron;
- The role of temporal integration and the period of inactivity following the excitatory period are usually ignored in the design of artificial neuron models;
- Artificial neurons themselves are non-time-varying, i.e., their synaptic time delays and synaptic strengths are constants.

Biological neurons in the M-P model



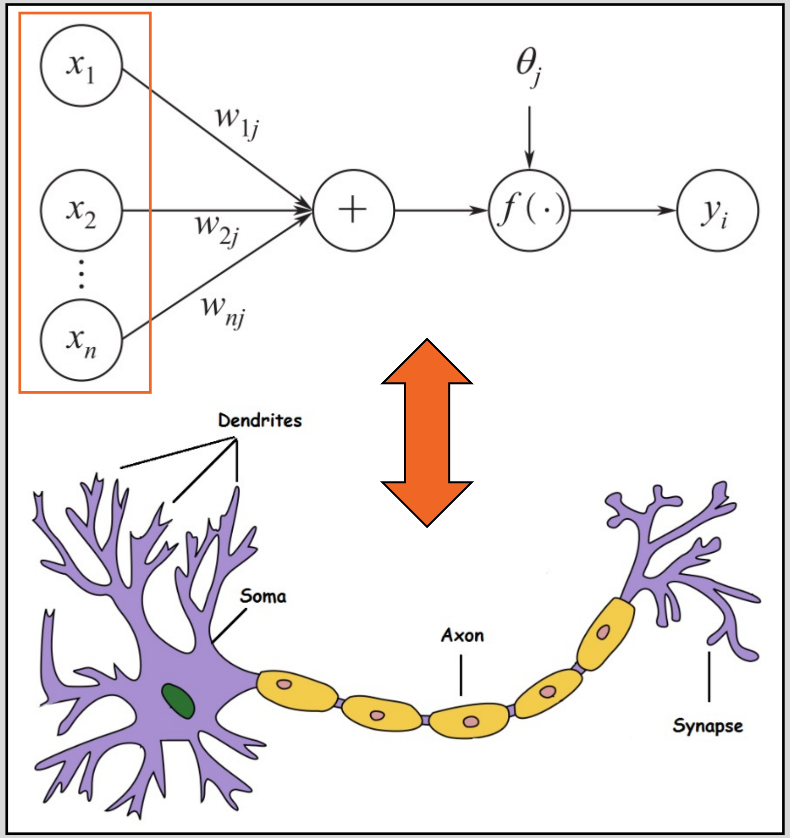
Biological Neuron	M-P Model
Neuron	j
Inputs	x_i
Weights	w_{ij}
Output	y_j
Summation	$\sum$
Membrane Potential	$\sum_{i=1}^{n} w_{ij}x_i$
Thresholds	θ

M-P Model-Neuron



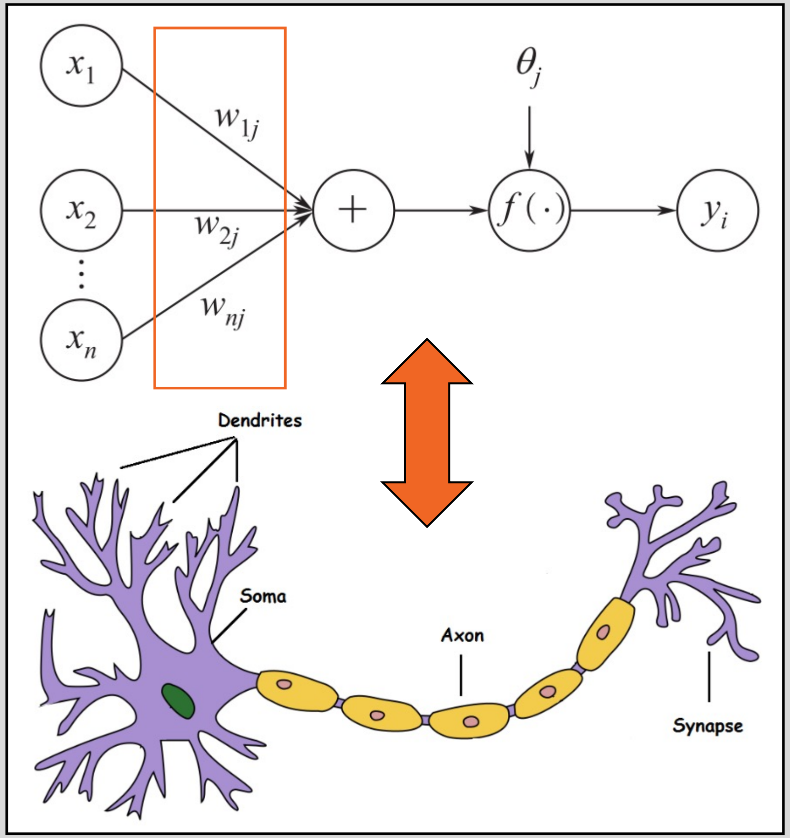
For a certain artificial neuron j (here j refers to a certain artificial neuron in the artificial neural network, which plays the role of an identifier), it is connected to multiple artificial neurons and can receive multiple input signals.

M-P Model-Inputs



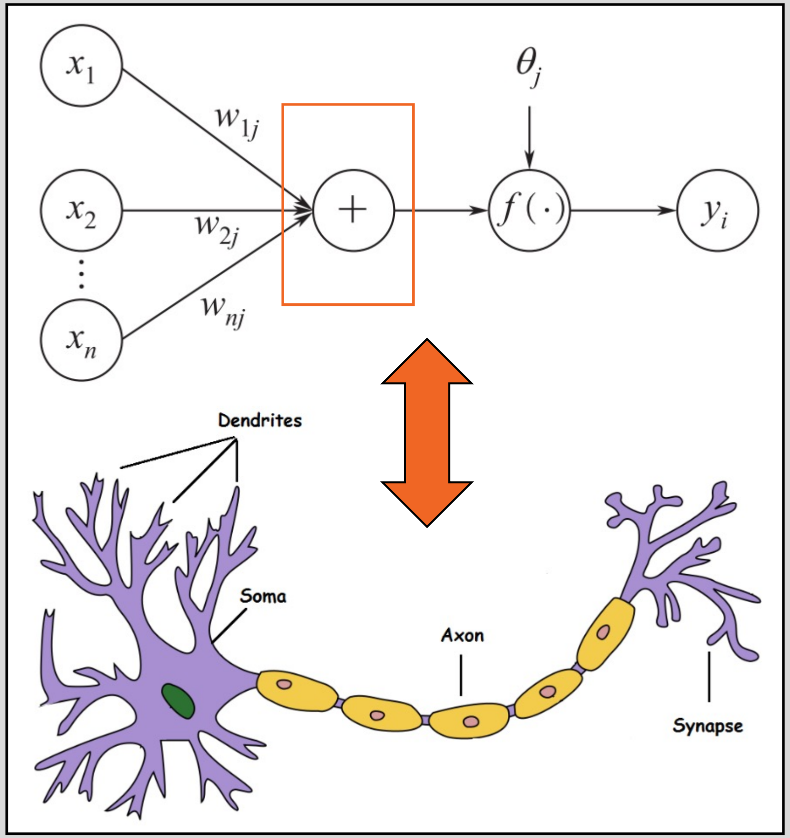
For a neuron can receive multiple input signals, denoted by $x_i (i = 1,2,\cdots,n)$.

M-P Model-Weights



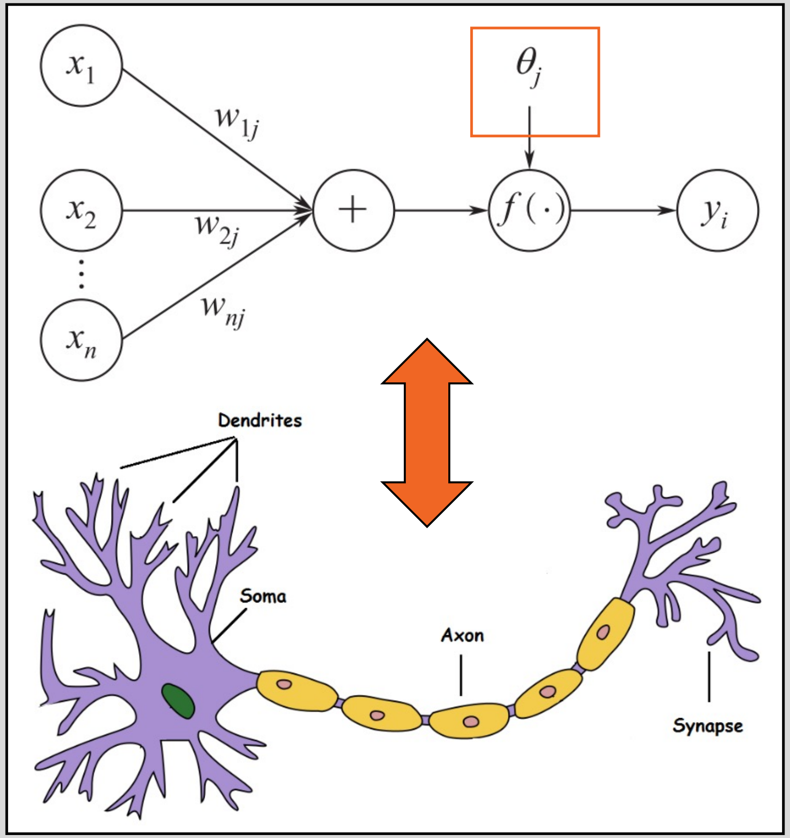
Given that biological neurons have different synaptic properties and synaptic strengths, they have different effects on biological neurons that receive signals. In artificial neurons we denote by weights w_{ij}, whose positivity and negativity simulate synaptic excitation and inhibition in biological neurons, and whose magnitude represents the different connection strengths of the synapses.

M-P Model-Summation



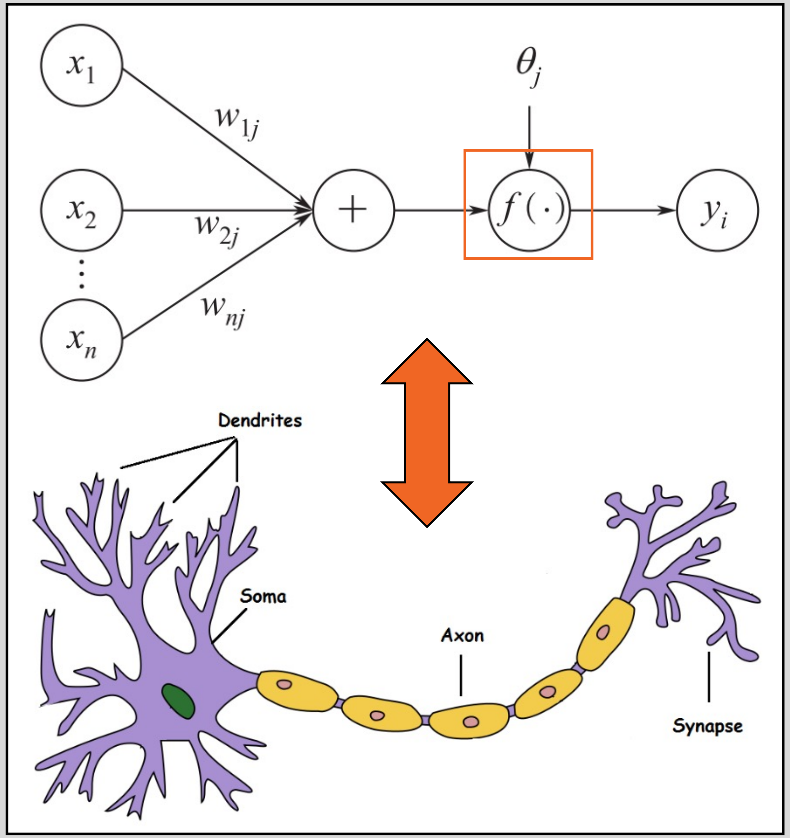
Similar to biological neurons, artificial neurons can cumulatively integrate input signals, equivalent to membrane potentials in biological neurons.

M-P Model-Thresholds



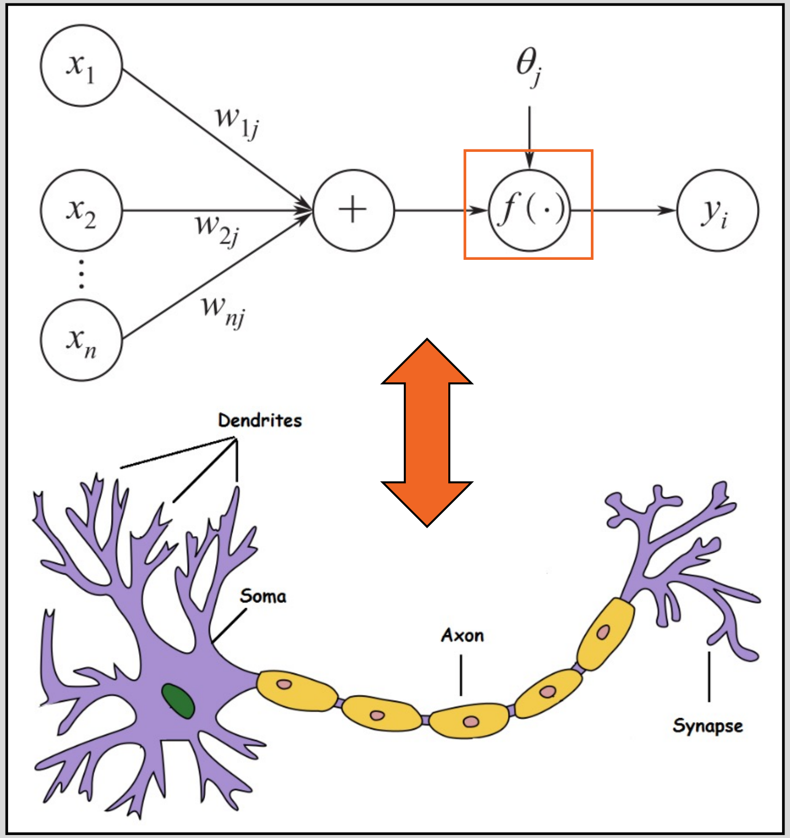
Whether a biological neuron is activated or not depends on a certain threshold level, i.e., a biological neuron is activated and thus releases an impulse signal only when the sum of inputs exceeds the threshold θ_j, otherwise the biological neuron does not occur to output an impulse signal.

M-P Model-Activation Function



The output of the final artificial neuron j is "0" or "1", and the function $f(\cdot)$ is called the **activation function**, **transfer function**, or **nonlinear function**.

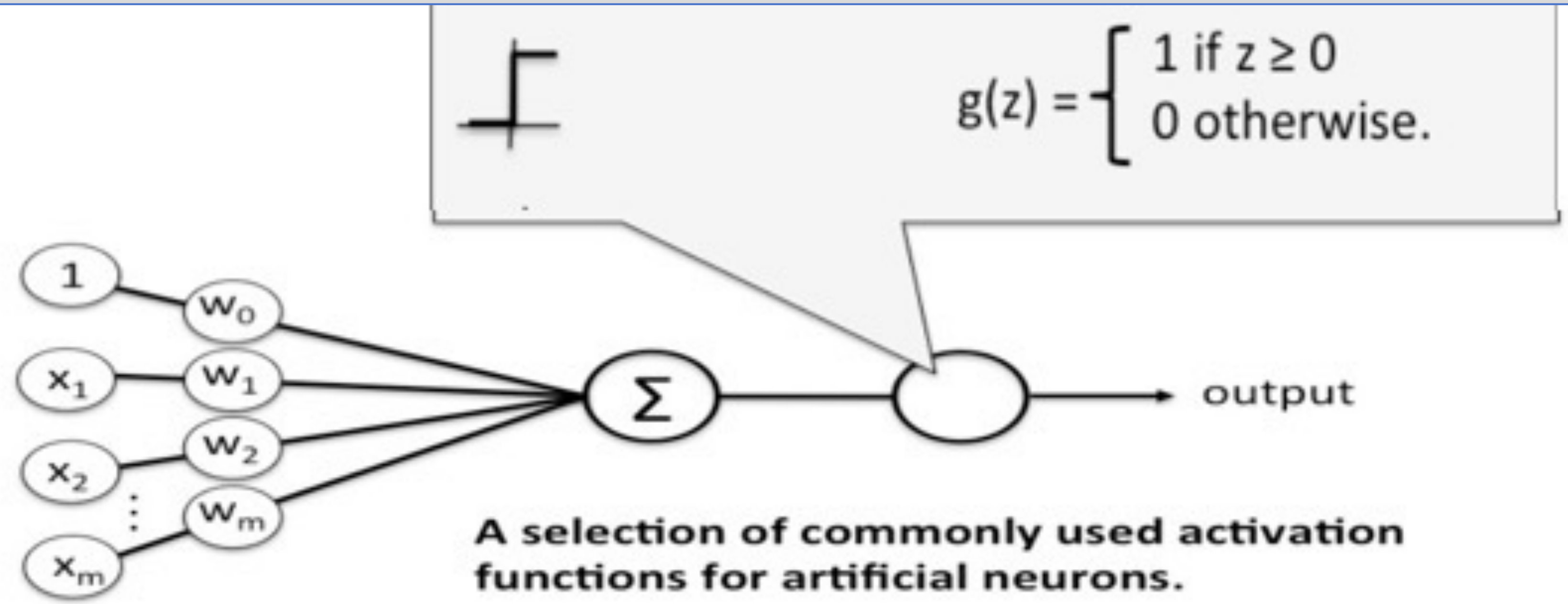
M-P Model-Output



The final artificial neuron output is

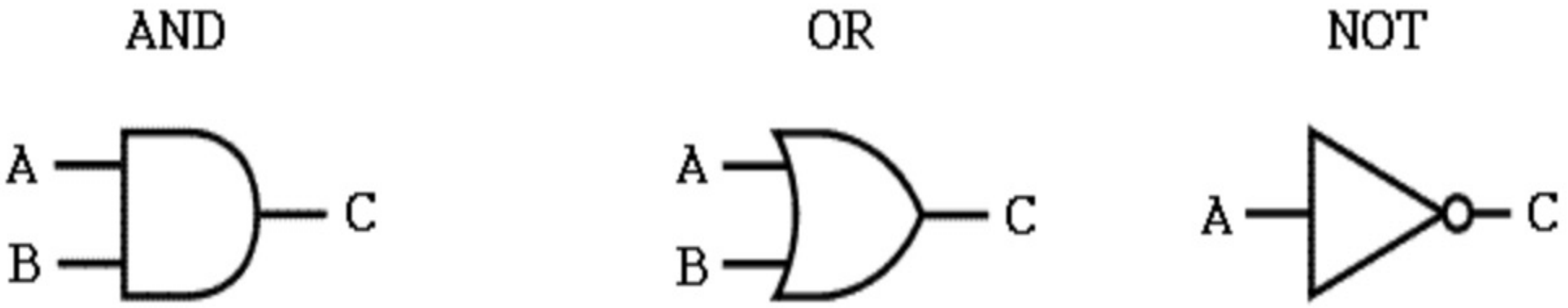
$$y_j = f(\sum_{i=1}^{n} w_{ij}x_i - \theta_j)$$

M-P Model



A selection of commonly used activation functions for artificial neurons.

Boolean function expressions for M-P models



AND

Inputs		Output
A	B	C
0	0	0
0	1	0
1	0	0
1	1	1

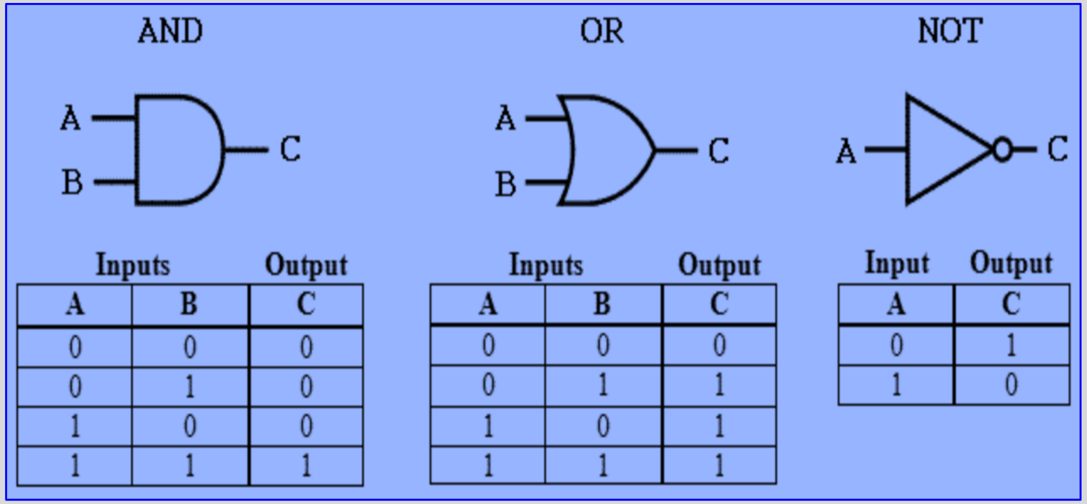
OR

Inputs		Output
A	B	C
0	0	0
0	1	1
1	0	1
1	1	1

NOT

Input	Output
A	C
0	1
1	0

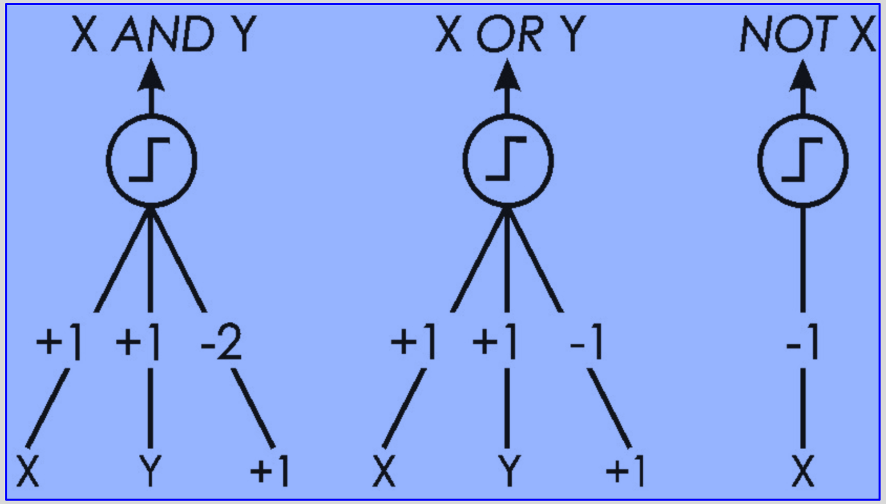
M-P Model and Logic

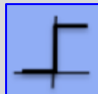


Inputs		Output
A	B	C
0	0	0
0	1	0
1	0	0
1	1	1

Inputs		Output
A	B	C
0	0	0
0	1	1
1	0	1
1	1	1

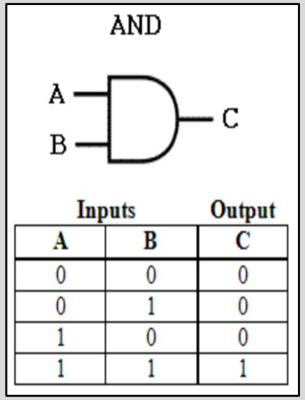
Input	Output
A	C
0	1
1	0

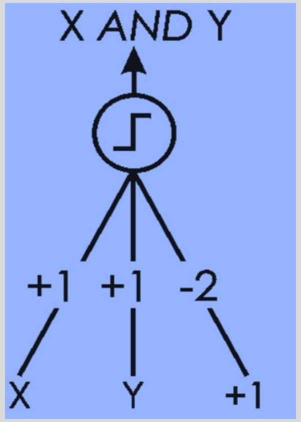




$$g(z) = \begin{cases} 1 \text{ if } z \geq 0 \\ 0 \text{ otherwise.} \end{cases}$$

M-P Model-AND



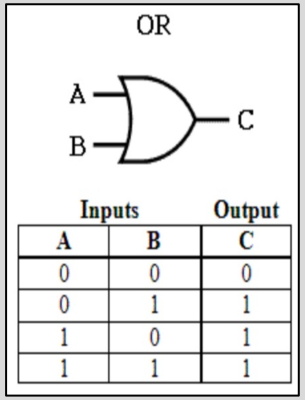


x_1	x_2	θ	w_1	w_2	$f(\cdot)$
0	0	2	1	1	$G(z)$
0	1	2	1	1	$G(z)$
1	0	2	1	1	$G(z)$
1	1	2	1	1	$G(z)$

$$G(z) = \begin{cases} 1 & if\ z \geq 0 \\ 0 & otherwise \end{cases}$$

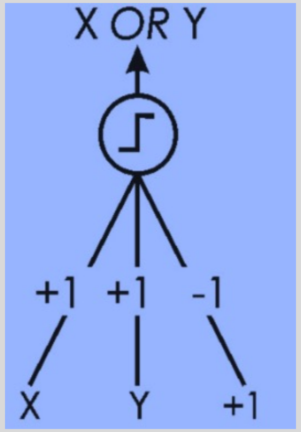
$z = w_1 x_1 + w_2 x_2 - \theta$	Output $y = \mathrm{G}(z)$	A AND B
0*1 + 0*1 – 2 = -2	0	0
0*1 + 1*1 – 2 = -1	0	0
1*1 + 0*1 – 2 = -1	0	0
1*1 + 1*1 –2= 0	1	1

M-P Model-OR



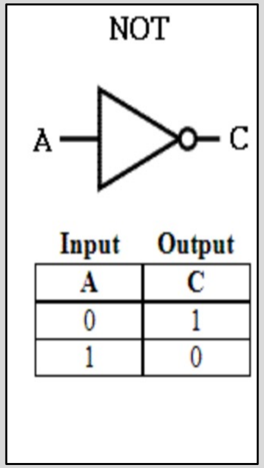
x_1	x_2	θ	w_1	w_2	$f(\cdot)$
0	0	1	1	1	$G(z)$
0	1	1	1	1	$G(z)$
1	0	1	1	1	$G(z)$
1	1	1	1	1	$G(z)$

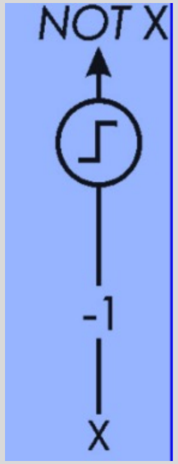
$$G(z) = \begin{cases} 1 & if\ z \geq 0 \\ 0 & otherwise \end{cases}$$



$z = w_1x_1 + w_2x_2 - \theta$	Output $y = \mathrm{G}(z)$	A OR B
0*1 + 0*1 – 1 = -1	0	0
0*1 + 1*1 – 1 = 0	1	1
1*1 + 0*1 – 1 = 0	1	1
1*1 + 1*1 – 1 = 1	1	1

M-P Model-NOT





x_1	θ	w_1	$f(\cdot)$
0	1	-1	$G(z)$
1	1	-1	$G(z)$

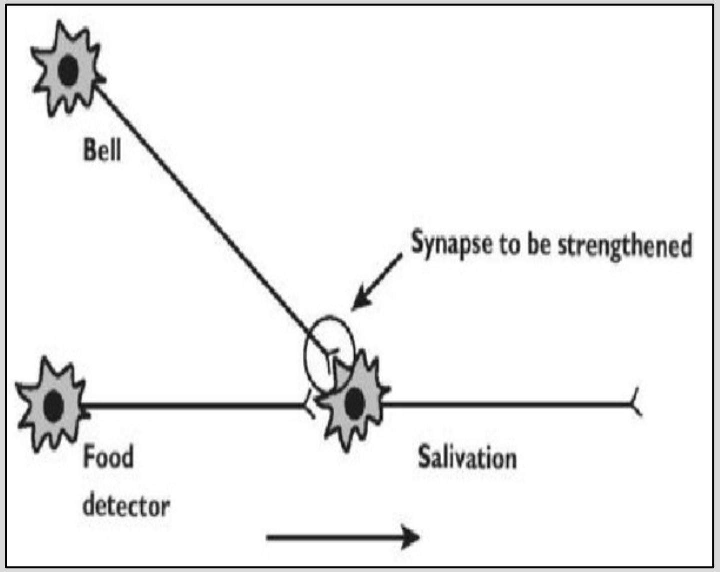
$$G(z) = \begin{cases} 1 & if\ z \geq 0 \\ 0 & otherwise \end{cases}$$

$z = w_1 x_1 - \theta$	Output $y = \mathrm{G}(z)$	NOT A
1 * (-1) - 0 = -1	0	0
0 *(-1) – 0 = 0	1	1

Q2:Are there any drawbacks to the M-P model ?

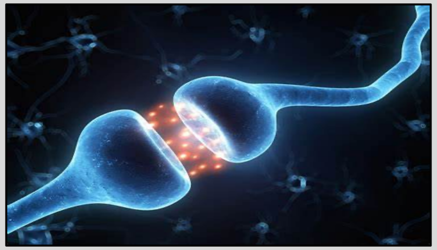
Lecture 5

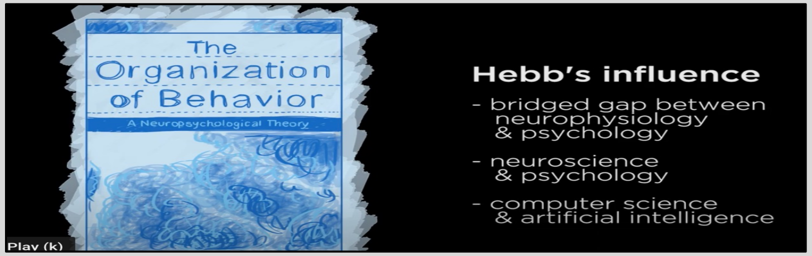
Hebb’ s Rule



When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A’ s efficiency, as one of the cells firing B, is increased.

Hebb





- Donald Olding Hebb (1904-1985) was a Canadian psychologist and a pioneer in neuropsychology (the study of the relationship between psychology and neuroscience).
- His book "The Organization of Behavior " (1949) proposed a theory of behavior based on the physiology of the nervous system. The book made an important attempt to find common ground between neurological and psychological concepts.

Hebb Learning Rule

$$w_{ij}(t+1) = w_{ij}(t) + \alpha y_j(t) y_i(t)$$

- Where α denotes the learning rate, $w_{ij}(t+1)$ and $w_{ij}(t)$ denote the connectivity of the artificial neuron j to the artificial neuron i at the moment of $t+1$ and t respectively strength, while y_i and y_j are the outputs of artificial neurons i and j. The Hebb learning rule belongs to the category of **unsupervised learning** algorithms, and the main idea is to adjust the connectivity relationship of two artificial neurons according to their excitation states, thus realizing the simple simulation of biological neural activity

Q3:Are there any disadvantages of Hebb learning rule

Delta Learning Rule

$$w_{ij}(t+1) = w_{ij}(t) + \alpha(d_i - y_i)y_j(t)$$

- Following the Hebb learning rule, the supervised Delta learning rule for artificial neurons was proposed to solve the problem of learning neuron weights when the inputs and outputs are known.

Delta Learning Rule

$$w_{ij}(t+1) = w_{ij}(t) + \alpha(d_i - y_i)y_j(t)$$

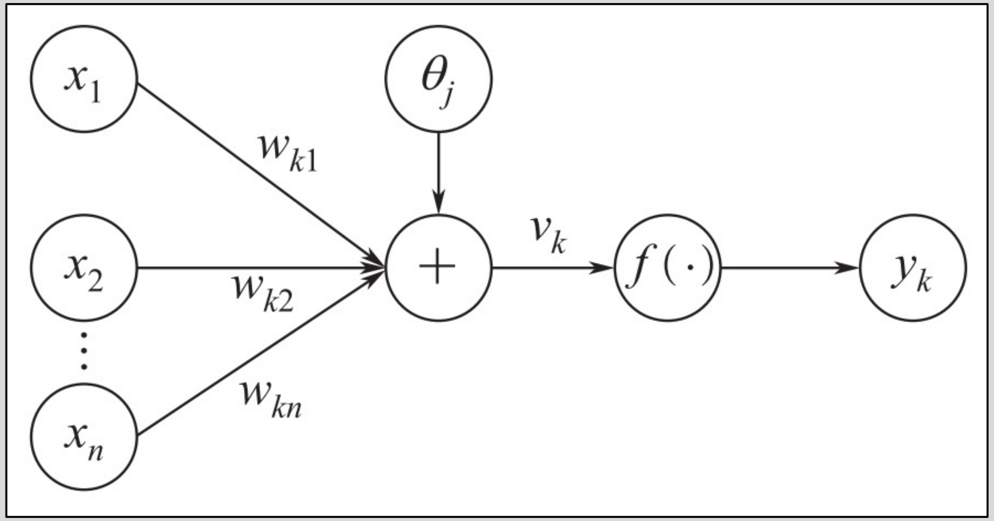
- where α denotes the learning rate, d_i and y_i are the desired and actual outputs of the artificial neuron i, and $y_j(t)$ denotes the state of the artificial neuron j at the moment t (activated or inhibited). Intuitively, when the actual output of artificial neuron i is larger than the desired output, the weight of the connection to the activated neuron is decreased while the weight of the connection to the inhibited neuron is increased; when the actual output of artificial neuron i is smaller than the desired output, the weight of the connection to the activated neuron is increased while the weight of the connection to the inhibited neuron is decreased.

Lecture 5

1. Lecture 4 Review
2. M-P Model
3. Hebb’ s Law and Delta Law
4. General Form of Artificial Neuron

Q4:Any other improvements to the M-P model?

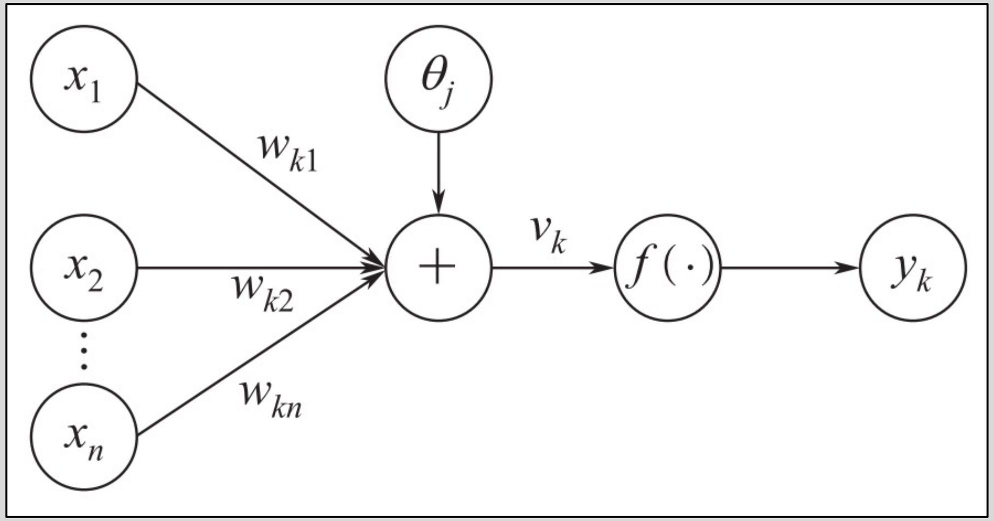
Mathematical Modeling of Artificial Neurons



A relatively standard mathematical model of a neuron, which consists of three main components:

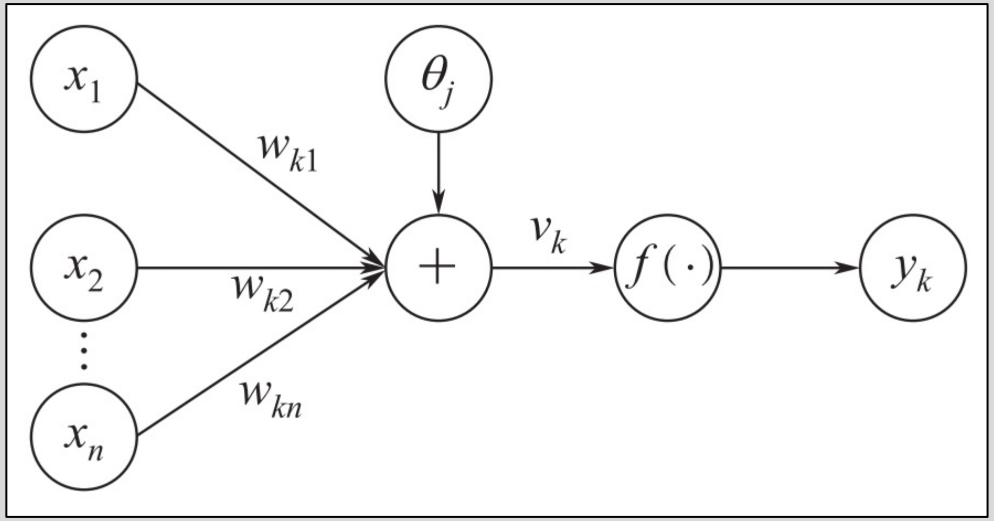
- **Weighting**
- **Linear summation**
- **Nonlinear function mapping**

Mathematical Modeling of Artificial Neurons -Weighting



Using $x_j\ (j = 1, \dots, n)$ represents the input of artificial neuron i, and w_{ji} denotes the value of the connection weights between the input artificial neuron node j and the artificial neuron node i, and a weighting operation of each input with its corresponding weights $x_{ji}x_j$ can be used to obtain u_i;

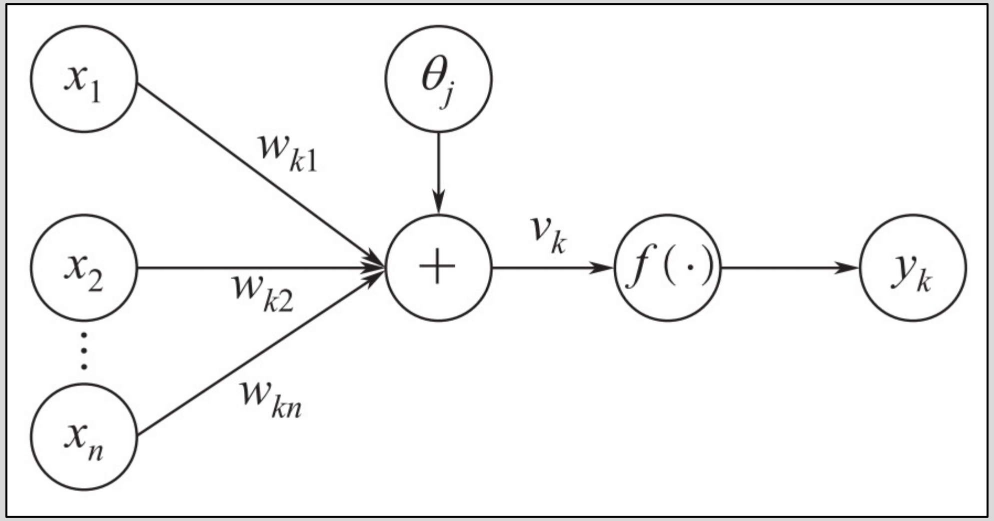
Mathematical Modeling of Artificial Neurons -Linear Summation



Summing over all μ_i $(i = 1, \ldots, n)$ and adding the bias θ_j , yields the adjusted weighted summation value v_k.

$$v_k = \sum_{i=1}^{n} w_{ki} x_j + \theta_k$$

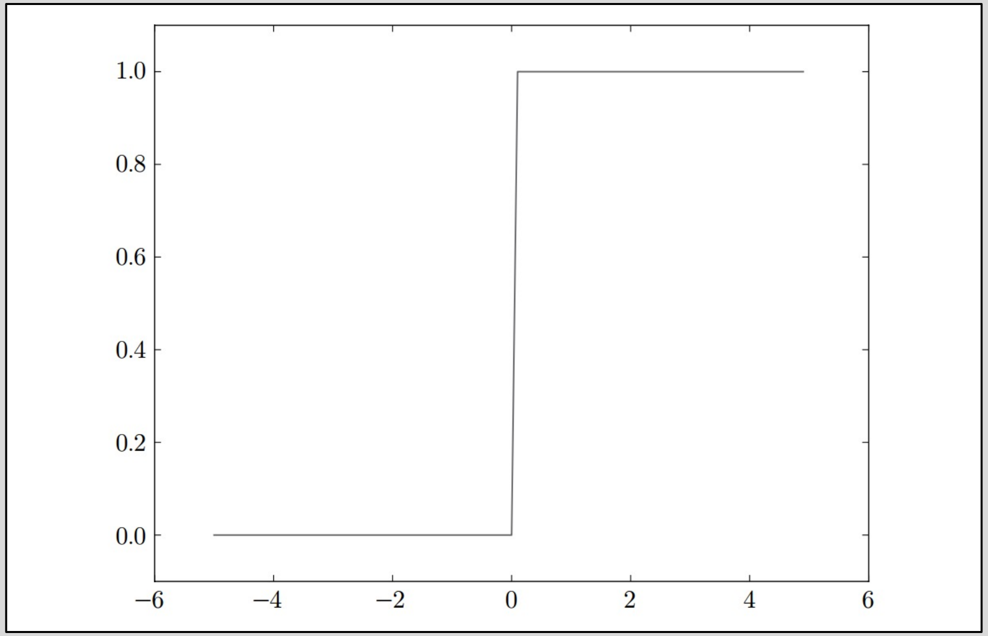
Mathematical Modeling of Artificial Neurons -Nonlinear Function Mapping



Also known as the activation function, its action simulates the phenomenon in which a biological neuron produces an excitatory signal after receiving a certain amount of stimulation, whereas if the stimulation is insufficient, the neuron remains in an inhibitory stat

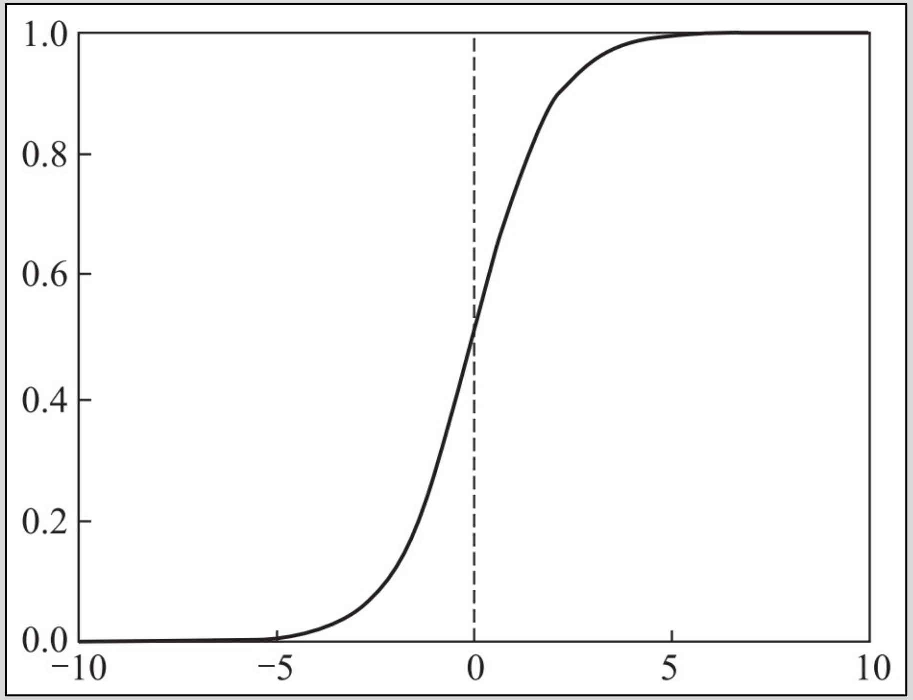
Q5:List some activation functions that you know?

Step Function



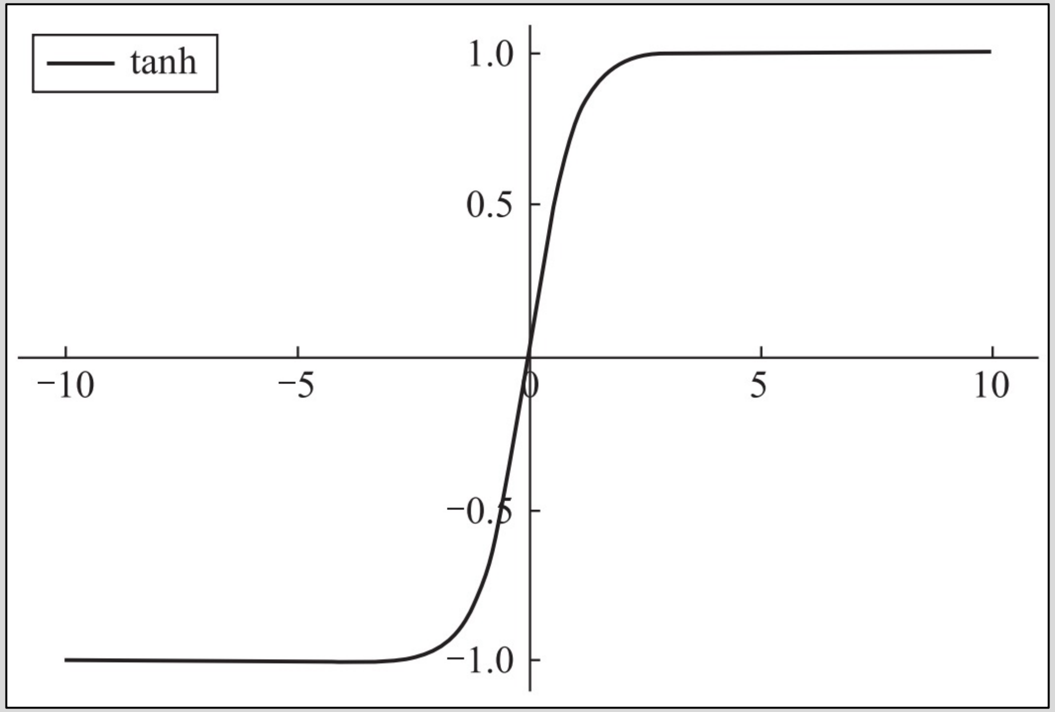
$$f(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases}$$

Sigmoid Function



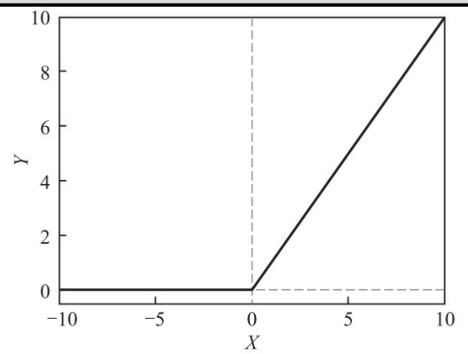
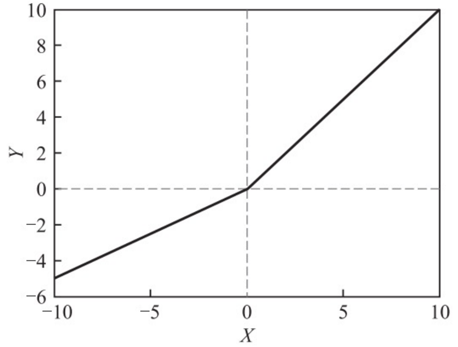
$$Sigmoid(x) = \frac{1}{1 + e^{-x}}$$

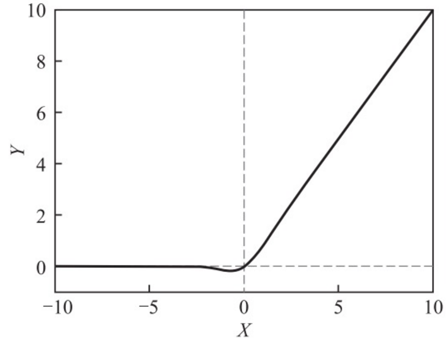
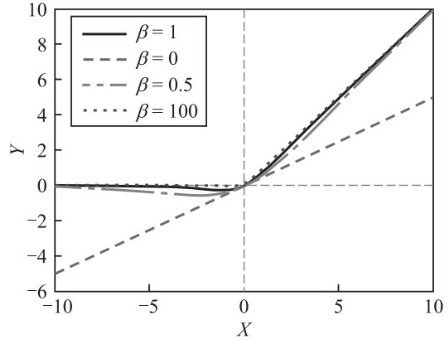
Tanh function



$$Tanh(x) = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

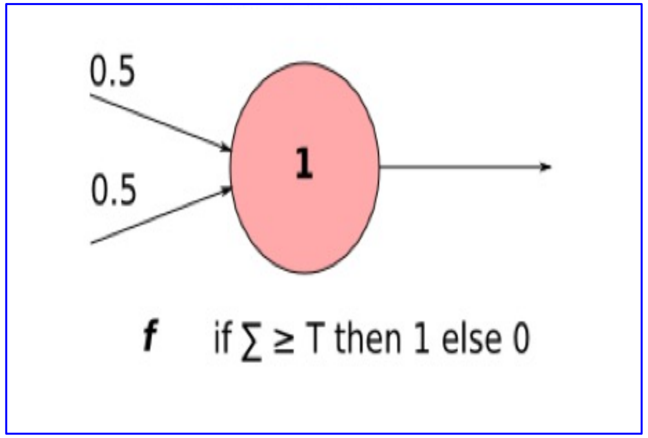
Other activation functions

$\mathrm{ReLU}(x)=\begin{cases}x, x \geqslant 0\\0, x<0\end{cases}=\max(0,x)$	
$\mathrm{LeakyReLU}(x)=\begin{cases}x, x>0\\\gamma x, x \leqslant 0\end{cases}$ $=\max(0,x)+\gamma\min(0,x)$	

$\mathrm{GELU}(x)=xP(X \leqslant x)$	
$\mathrm{swish}(x)=x\sigma(\beta x)$	

Q6 : In Which Areas, General Artificial Neuron is inferior to Biological Neuron and In Which Areas superior ?

Homework1 : Prove the M-P Model is an “AND” Logic



Inputs		Output
A	B	C
0	0	0
0	1	0
1	0	0
1	1	1

Transfer Function is G (SUM-T)

Homework 2

- How AI Can Help Education?
- How Shall We Integrate AI in CS 103 Module?

Introduction to AI （CS103） – 05

The Basis of AI - Artificial Neurons

Jimmy Liu 刘江

2023-10-20